

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: LI, ET AL.
- (ii) TITLE OF INVENTION: Connective Tissue Growth Factor-2
- (iii) NUMBER OF SEQUENCES: 2
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
- (B) STREET: 6 BECKER FARM ROAD
- (C) CITY: ROSELAND
- (D) STATE: NEW JERSEY
- (E) COUNTRY: USA
- (F) ZIP: 07068
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5 INCH DISKETTE
- (B) COMPUTER: IBM PS/2
- (C) OPERATING SYSTEM: MS-DOS
- (D) SOFTWARE: WORD PERFECT 5.1
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE: Concurrently
- (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA
- (A) APPLICATION NUMBER: PCT/US94/07736
- (B) FILING DATE: 12 JUL 94

(viii) ATTORNEY/AGENT INFORMATION:
 (A) NAME: FERRARO, GREGORY D.
 (B) REGISTRATION NUMBER: 36,134
 (C) REFERENCE/DOCKET NUMBER: 325800-317

(ix) TELECOMMUNICATION INFORMATION:
 (A) TELEPHONE: 201-994-1700
 (B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 1128 BASE PAIRS
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: SINGLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGAGCTCCC	GAATCGTCAG	GGAGCTCGCC	TTAGTCGTCA	CCCTTCTCCA	CTTGACCAGG	60
GTGGGGCTCT	CCACCTGCCC	CGCTGACTGC	CACTGCCCCC	TGGAGGCGCC	CAAGTGCGCG	120
CCGGGAGTCG	GGCTGTCCG	GGACGGCTGC	GGCTGTTGTA	AGGTCTGCGC	CAAGCAGCTC	180
AACGAGGACT	GCAG/AAAAC	GCAGCCCTGC	GACCACACCA	AGGGGCTGGA	ATGCAACTTC	240
GGCGCCAGCT	CCACCGCTCT	GAAGGGGATC	TGCAGAGCTC	AGTCAGAGGG	CAGACCCTGT	300
GAATATAACT	CCAG/ATCTA	CCAAAACGGG	GAAAGTTTCC	AGCCCAACTG	TAAACATCAG	360
TGCACATGTA	TTGG/ATGGCG	CCGGGGGGCT	TGCATTCTTC	TGTGTCCCCA	AGAACTATCT	420
CCCCCAACT	TGGG/TGTCC	CAACCTCTCG	CTGGTCAAAG	TTACCGGGCA	GTGCTGCGAG	480
GAGTGGGTCT	GTGA/GAGGA	TAGTATCAAG	GACCCCATGG	AGGACCAGGA	CGGCCTCCTT	540
GGCAAGGGGC	TGGG/ATTGA	TGCCTCCGAG	GTGGAGTTGA	CGAGAAACAA	TGAATTGATT	600
GCAGTTGGAA	AAGG/TAGCTC	ACTGAAGCGG	CTCCCTGTTT	TTGGAATGGA	GCCTCGCATC	660
CTATACAACC	CTTT/CAAGG	CCAGAAATGT	ATTGTTCAAA	CAACTTCATG	GTCCCAAGTGC	720
TCAAAGACCT	GTG/AACTGG	TATCTCCACA	CGAGTTACCA	ATGACAACCC	TGAGTGCCGC	780
CTTGTGAAAG	AAACCCGGAT	TTGTGAGGTG	CGGCCTTGTC	GACAGCCAGT	GTACAGCAGC	840
CTGAAAAGG	GCAAGAAATG	CAGCAAGACC	AAGAAATCCC	CCGAACCACT	CAGGTTTACT	900
TACGCTGGAT	GTTTGAGTGT	GAAGAAATAC	CGGCCCAAGT	ACTGCGGTTC	CTGCGTGGAC	960

GGCCGATGCT GCACGCTCCA GCTGACCAGG ACTGTGAAGA TCGCGTTCCC CTGCGAAGAT
 GGGGAGACAT TTTCCAAGAA CGTCATGATG ATCCAGTCCT CCAAATGCAA CTACAACTGC
 CCGCATGCCA ARGAAAGAGC GTTTCCTTC TACAGGCTGT TCCAATGA

1020

1080

1128

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 375 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS:

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ser	Ser	Arg	Ile	Val	Arg	Glu	Leu	Ala	Leu	Val	Val	Thr	Leu	
				-20					-15					-10	
Leu	His	Leu	Thr	Arg	Val	Gly	Leu	Ser	Thr	Cys	Pro	Ala	Asp	Cys	
				-5					1					5	
His	Cys	Pro	Leu	Glu	Ala	Pro	Lys	Cys	Ala	Pro	Gly	Val	Gly	Leu	
				10					15					20	
Val	Arg	Asp	Gly	Cys	Gly	Cys	Cys	Lys	Val	Cys	Ala	Lys	Gln	Leu	
				25					30					35	
Asn	Glu	Asp	Cys	Arg	Lys	Thr	Gln	Pro	Cys	Asp	His	Thr	Lys	Gly	
				40					45					50	
Leu	Glu	Cys	Asn	Phe	Gly	Ala	Ser	Ser	Thr	Ala	Leu	Lys	Gly	Ile	
				55					60					65	
Cys	Arg	Ala	Gln	Ser	Glu	Gly	Arg	Pro	Cys	Glu	Tyr	Asn	Ser	Arg	
				70					75					80	
Ile	Tyr	Gln	Asn	Gly	Glu	Ser	Phe	Gln	Pro	Asn	Cys	Lys	His	Gln	
				85					90					95	
Cys	Thr	Cys	Ile	Gly	Thr	Arg	Arg	Gly	Ala	Cys	Ile	Pro	Leu	Cys	
				100					105					110	
Pro	Gln	Glu	Leu	Ser	Leu	Pro	Asn	Leu	Gly	Cys	Pro	Asn	Pro	Arg	
				115					120					125	

Leu	Val	Lys	Val	Thr	Gly	Gln	Cys	Cys	Glu	Glu	Trp	Val	Cys	Asp	130	135	140
Glu	Asp	Ser	Ile	Lys	Asp	Pro	Met	Glu	Asp	Gln	Asp	Gly	Leu	Leu	145	150	155
Gly	Lys	Gly	Leu	Gly	Phe	Asp	Ala	Ser	Glu	Val	Glu	Leu	Thr	Arg	160	165	170
Asn	Asn	Glu	Leu	Ile	Ala	Val	Gly	Lys	Gly	Ser	Ser	Leu	Lys	Arg	175	180	185
Leu	Pro	Val	Phe	Gly	Met	Glu	Pro	Arg	Ile	Leu	Tyr	Asn	Pro	Leu	190	195	200
Gln	Gly	Gln	Lys	Cys	Ile	Val	Gln	Thr	Thr	Ser	Trp	Ser	Gln	Cys	205	210	215
Ser	Lys	Thr	Cys	Gly	Thr	Gly	Ile	Ser	Thr	Arg	Val	Thr	Asn	Asp	220	225	230
Asn	Pro	Glu	Cys	Arg	Leu	Val	Lys	Glu	Thr	Arg	Ile	Cys	Gly	Val	235	240	245
Arg	Pro	Cys	Gly	Gln	Pro	Val	Tyr	Ser	Ser	Leu	Lys	Lys	Gly	Lys	250	255	260
Lys	Cys	Ser	Lys	Thr	Lys	Lys	Ser	Pro	Glu	Pro	Val	Arg	Phe	Thr	265	270	275
Tyr	Ala	Gly	Cys	Leu	Ser	Val	Lys	Lys	Tyr	Arg	Pro	Lys	Tyr	Cys	280	285	290
Gly	Ser	Cys	Val	Asp	Gly	Arg	Cys	Cys	Thr	Pro	Gln	Leu	Thr	Arg	295	300	305
Thr	Val	Lys	Met	Arg	Phe	Pro	Cys	Glu	Asp	Gly	Glu	Thr	Phe	Ser	310	315	320
Lys	Asn	Val	Met	Met	Ile	Gln	Ser	Ser	Lys	Cys	Asn	Tyr	Asn	Cys	325	330	335
Pro	His	Ala	Asn	Glu	Ala	Ala	Phe	Pro	Phe	Tyr	Arg	Leu	Phe	Gln	340	345	350

38
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: LI, HAODONG
ADAMS, MARK D

(ii) TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2

(iii) NUMBER OF SEQUENCES: 8

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
(B) STREET: 9410 KEY WEST AVENUE
(C) CITY: ROCKVILLE
(D) STATE: MD
(E) COUNTRY: US
(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/459,101
(B) FILING DATE: 02-JUN-1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: A. ANDERS BROOKES
(B) REGISTRATION NUMBER: 36,373
(C) REFERENCE/DOCKET NUMBER: PF126P1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 301-309-8504
(B) TELEFAX: 301-309-8439

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1127 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..1122

SUB
B

003441-070000

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG AGC TCC CGA ATC GTC AGG GAG CTC GCC TTA GTC GTC ACC CTT CTC 48
 Met Ser Ser Arg Ile Val Arg Glu Leu Ala Leu Val Val Thr Leu Leu
 1 5 10 15
 CAC TTG ACC AGC GTG GGG CTC TCC ACC TGC CCC GCT GAC TGC CAC TGC 96
 His Leu Thr Arg Val Gly Leu Ser Thr Cys Pro Ala Asp Cys His Cys
 20 25 30
 CCC CTG GAG GCG CCC AAG TGC GCG CCG GGA GTC GGG CTG GTC CGG GAC 144
 Pro Leu Glu Ala Pro Lys Cys Ala Pro Gly Val Gly Leu Val Arg Asp
 35 40 45
 GGC TGC GGC TGT TGT AAG GTC TGC GCC AAG CAG CTC AAC GAG GAC TGC 192
 Gly Cys Gly Cys Cys Lys Val Cys Ala Lys Gln Leu Asn Glu Asp Cys
 50 55 60
 AGA AAA ACG CAG CCC TGC GAC CAC ACC AAG GGG CTG GAA TGC AAC TTC 240
 Arg Lys Thr Gln Pro Cys Asp His Thr Lys Gly Leu Glu Cys Asn Phe
 65 70 75 80
 GGC GCC AGC TCC ACC GCT CTG AAG GGG ATC TGC AGA GCT CAG TCA GAG 288
 Gly Ala Ser Ser Thr Ala Leu Lys Gly Ile Cys Arg Ala Gln Ser Glu
 85 90 95
 GGC AGA CCC TGT GAA TAT AAC TCC AGA ATC TAC CAA AAC GGG GAA AGT 336
 Gly Arg Pro Cys Glu Tyr Asn Ser Arg Ile Tyr Gln Asn Gly Glu Ser
 100 105 110
 TTC CAG CCC AAC TGT AAA CAT CAG TGC ACA TCT ATT GGA TGG CGC CGG 384
 Phe Gln Pro Asn Cys Lys His Gln Cys Thr Cys Ile Gly Trp Arg Arg
 115 120 125
 GGC GCT TGC ATT CCT CTG TGT CCC CAA GAA CTA TCT CTC CCC AAC TTG 432
 Gly Ala Cys Ile Pro Leu Cys Pro Gln Glu Leu Ser Leu Pro Asn Leu
 130 135 140
 GGC TGT CCC AAC CCT CGG CTG GTC AAA GTT ACC GGG CAG TGC TGC GAG 480
 Gly Cys Pro Asn Pro Arg Leu Val Lys Val Thr Gly Gln Cys Cys Glu
 145 150 155 160
 GAG TGG GTC TGT GAC GAG GAT AGT ATC AAG GAC CCC ATG GAG GAC CAG 528
 Glu Trp Val Cys Asp Glu Asp Ser Ile Lys Asp Pro Met Glu Asp Gln
 165 170 175
 GAC GGC CTC CTT GGC AAG GGG CTG GGA TTC GAT GCC TCC GAG GTG GAG 576
 Asp Gly Leu Leu Gly Lys Gly Leu Gly Phe Asp Ala Ser Glu Val Glu
 180 185 190
 TTG ACG AGA AAC AAT GAA TTG ATT GCA GTT GGA AAA GGC AGC TCA CTG 624
 Leu Thr Arg Asn Asn Glu Leu Ile Ala Val Gly Lys Gly Ser Ser Leu
 195 200 205

51304650
 SUB
 B1

AAG CGG CTC CCT GTT TTT GGA ATG GAG CCT CGC ATC CTA TAC AAC CCT 672
 Lys Arg Leu Pro Val Phe Gly Met Glu Pro Arg Ile Leu Tyr Asn Pro
 210 215 220

 TTA CAA GGC CAG AAA TGT ATT GTT CAA ACA ACT TCA TGG TCC CAG TGC 720
 Leu Gln Gly Gln Lys Cys Ile Val Gln Thr Thr Ser Trp Ser Gln Cys
 225 230 235 240

 TCA AAG ACC TGT GGA ACT GGT ATC TCC ACA CGA GTT ACC AAT GAC AAC 768
 Ser Lys Thr Cys Gly Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn
 245 250 255

 CCT GAG TGC CGC CTT GTG AAA GAA ACC CGG ATT TGT GAG GTG CGG CCT 816
 Pro Glu Cys Arg Leu Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro
 260 265 270

 TGT GGA CAG CCA GTG TAC AGC AGC CTG AAA AAG GGC AAG AAA TGC AGC 864
 Cys Gly Gln Pro Val Tyr Ser Ser Leu Lys Lys Gly Lys Lys Cys Ser
 275 280 285

 AAG ACC AAG AAA TCC CCC GAA CCA GTC AGG TTT ACT TAC GCT GGA TGT 912
 Lys Thr Lys Lys Ser Pro Glu Pro Val Arg Phe Thr Tyr Ala Gly Cys
 290 295 300

 TTG AGT GTG AAG AAA TAC CGG CCC AAG TAC TGC GGT TCC TGC GTG GAC 960
 Leu Ser Val Lys Lys Tyr Arg Pro Lys Tyr Cys Gly Ser Cys Val Asp
 305 310 315 320

 GGC CGA TGC TGC ACG CCC CAG CTG ACC AAG ACT GTG AAG ATG CGG TTC 1008
 Gly Arg Cys Cys Thr Pro Gln Leu Thr Arg Thr Val Lys Met Arg Phe
 325 330 335

 CCC TGC GAA GAT GGG GAG ACA TTT TCC AAG AAC GTC ATG ATG ATC CAG 1056
 Pro Cys Glu Asp Gly Glu Thr Phe Ser Lys Asn Val Met Met Ile Gln
 340 345 350

 TCC TCC AAA TGC AAC TAC AAC TGC CCG CAT GCC AAG AAG CAG CGT TTC 1104
 Ser Ser Lys Cys Asn Tyr Asn Cys Pro His Ala Lys Lys Gln Arg Phe
 355 360 365

 CCT TCT ACA GGC TGT TCC AATGA 1127
 Pro Ser Thr Gly Cys Ser
 370

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Ser Arg Ile Val Arg Glu Leu Ala Leu Val Val Thr Leu Leu
 1 5 10 15
 His Leu Thr Arg Val Gly Leu Ser Thr Cys Pro Ala Asp Cys His Cys
 20 25 30
 Pro Leu Glu Ala Pro Lys Cys Ala Pro Gly Val Gly Leu Val Arg Asp
 35 40 45
 Gly Cys Gly Cys Cys Lys Val Cys Ala Lys Gln Leu Asn Glu Asp Cys
 50 55 60
 Arg Lys Thr Gln Pro Cys Asp His Thr Lys Gly Leu Glu Cys Asn Phe
 65 70 75 80
 Gly Ala Ser Ser Thr Ala Leu Lys Gly Ile Cys Arg Ala Gln Ser Glu
 85 90 95
 Gly Arg Pro Cys Glu Tyr Asn Ser Arg Ile Tyr Gln Asn Gly Glu Ser
 100 105 110
 Phe Gln Pro Asn Cys Lys His Gln Cys Thr Cys Ile Gly Trp Arg Arg
 115 120 125
 Gly Ala Cys Ile Pro Leu Cys Pro Gln Glu Leu Ser Leu Pro Asn Leu
 130 135 140
 Gly Cys Pro Asn Pro Arg Leu Val Lys Val Thr Gly Gln Cys Cys Glu
 145 150 155 160
 Glu Trp Val Cys Asp Glu Asp Ser Ile Lys Asp Pro Met Glu Asp Gln
 165 170 175
 Asp Gly Leu Leu Gly Lys Gly Leu Gly Phe Asp Ala Ser Glu Val Glu
 180 185 190
 Leu Thr Arg Asn Asn Glu Leu Ile Ala Val Gly Lys Gly Ser Ser Leu
 195 200 205
 Lys Arg Leu Pro Val Phe Gly Met Glu Pro Arg Ile Leu Tyr Asn Pro
 210 215 220
 Leu Gln Gly Gln Lys Cys Ile Val Gln Thr Thr Ser Trp Ser Gln Cys
 225 230 235 240
 Ser Lys Thr Cys Gly Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn
 245 250 255
 Pro Glu Cys Arg Leu Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro
 260 265 270
 Cys Gly Gln Pro Val Tyr Ser Ser Leu Lys Lys Gly Lys Lys Cys Ser
 275 280 285

SUB B
662020-51884600

Pro Ser Thr Gly Cys Ser
370

Figure 1 consists of 12 sub-diagrams labeled (a) through (l), showing the sequential assembly of a mechanical linkage.
 (a) A single horizontal link.
 (b) A second link added at one end, forming a V-shape.
 (c) A third link added to the second, forming a Z-shape.
 (d) A fourth link added to the third, forming a more complex zig-zag.
 (e) A fifth link added, creating a closed loop with one fixed pivot.
 (f) A sixth link added, extending from the loop.
 (g) A seventh link added, further extending the chain.
 (h) An eighth link added, creating a second closed loop.
 (i) A ninth link added, extending from the second loop.
 (j) A tenth link added, further extending the chain.
 (k) An eleventh link added, creating a third closed loop.
 (l) The final twelfth link added, completing the 12-bar mechanism with two degrees of freedom.

B

- (ii) MOLECULE TYPE: protein

10-12-1962

Gly Arg Pro Cys Glu Tyr Asn Ser Arg Ile Tyr Gln Asn Gly Glu Ser
100 105 110

Phe Gln Pro Asn Cys Lys His Gln Cys Thr Cys Ile Asp Gly Ala Val
 115 120 125
 Gly Cys Ile Pro Leu Cys Pro Gln Glu Leu Ser Leu Pro Asn Leu Gly
 130 135 140
 Cys Pro Asn Pro Arg Leu Val Lys Val Ser Gly Gln Cys Cys Glu Glu
 145 150 155 160
 Trp Val Cys Asp Glu Asp Ser Ile Lys Asp Ser Leu Asp Asp Gln Asp
 165 170 175
 Asp Leu Leu Gly Leu Asp Ala Ser Glu Val Glu Leu Thr Arg Asn Asn
 180 185 190
 Glu Leu Ile Ala Ile Gly Lys Gly Ser Ser Leu Lys Arg Leu Pro Val
 195 200 205
 Phe Gly Thr Glu Pro Arg Val Leu Phe Asn Pro Leu His Ala His Gly
 210 215 220
 Gln Lys Cys Ile Val Gln Thr Thr Ser Trp Ser Gln Cys Ser Lys Ser
 225 230 235 240
 Cys Gly Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Pro Glu Cys
 245 250 255
 Arg Leu Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro Cys Gly Gln
 260 265 270
 Pro Val Tyr Ser Ser Leu Lys Lys Gly Lys Lys Cys Ser Lys Thr Lys
 275 280 285
 Lys Ser Pro Glu Pro Val Arg Phe Thr Tyr Ala Gly Cys Ser Ser Val
 290 295 300
 Lys Lys Tyr Arg Pro Lys Tyr Cys Gly Ser Cys Val Asp Gly Arg Cys
 305 310 315 320
 Cys Thr Pro Leu Gln Thr Arg Thr Val Lys Met Arg Phe Arg Cys Glu
 325 330 335
 Asp Gly Glu Met Phe Ser Lys Asn Val Met Met Ile Gln Ser Cys Lys
 340 345 350
 Cys Asn Tyr Asn Cys Pro His Pro Asn Glu Ala Ser Phe Arg Leu Tyr
 355 360 365
 Ser Leu Phe Asn
 370

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 345 amino acids

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Leu	Ala	Ser	Val	Ala	Gly	Pro	Ile	Ser	Leu	Ala	Leu	Val	Leu	Leu
1				5					10					15	
Ala	Leu	Cys	Thr	Arg	Pro	Ala	Thr	Gly	Gln	Asp	Cys	Ser	Ala	Gln	Cys
			20					25					30		
Gln	Cys	Ala	Ala	Glu	Ala	Ala	Pro	His	Cys	Pro	Ala	Gly	Val	Ser	Leu
		35					40					45			
Val	Leu	Asp	Gly	Cys	Gly	Cys	Cys	Arg	Val	Cys	Ala	Lys	Gln	Leu	Gly
	50					55					60				
Glu	Leu	Cys	Thr	Glu	Arg	Asp	Pro	Cys	Asp	Pro	His	Lys	Gly	Leu	Phe
65					70					75					80
Cys	Asp	Phe	Gly	Ser	Pro	Ala	Asn	Arg	Lys	Ile	Gly	Val	Cys	Thr	Ala
				85					90					95	
Lys	Asp	Gly	Ala	Pro	Cys	Val	Phe	Gly	Gly	Ser	Val	Tyr	Arg	Ser	Gly
			100					105					110		
Glu	Ser	Phe	Gln	Ser	Ser	Cys	Lys	Tyr	Gln	Cys	Thr	Cys	Leu	Asp	Gly
		115					120					125			
Ala	Val	Gly	Cys	Val	Pro	Leu	Cys	Ser	Met	Asp	Val	Arg	Leu	Pro	Ser
		130				135					140				
Pro	Asp	Cys	Pro	Phe	Pro	Arg	Arg	Val	Lys	Leu	Pro	Gly	Lys	Cys	Cys
145					150					155					160
Lys	Glu	Trp	Val	Cys	Asp	Glu	Pro	Lys	Asp	Arg	Thr	Ala	Val	Gly	Pro
				165					170					175	
Ala	Leu	Ala	Ala	Tyr	Arg	Leu	Glu	Asp	Thr	Phe	Gly	Pro	Asp	Pro	Thr
			180					185					190		
Met	Met	Arg	Ala	Asn	Cys	Leu	Val	Gln	Thr	Thr	Glu	Trp	Ser	Ala	Cys
		195					200					205			
Ser	Lys	Thr	Cys	Cys	Met	Gly	Ile	Ser	Thr	Arg	Val	Thr	Asn	Asp	Asn
		210				215					220				
Thr	Phe	Cys	Arg	Leu	Glu	Lys	Gln	Ser	Arg	Leu	Cys	Met	Val	Arg	Pro
225					230					235					240

Cys Glu Ala Asp Leu Glu Glu Asn Ile Lys Lys Gly Lys Lys Cys Ile
245 250 255

Arg Thr Pro Lys Ile Ala Lys Pro Val Lys Phe Glu Leu Ser Gly Cys
260 265 270

Thr Ser Val Lys Thr Tyr Arg Ala Lys Phe Cys Gly Val Cys Thr Asp
275 280 285

Gly Arg Cys Cys Thr Pro His Arg Thr Thr Thr Leu Pro Val Glu Phe
290 295 300

Lys Cys Pro Asp Gly Glu Ile Met Lys Lys Asn Met Met Phe Ile Lys
305 310 315 320

Thr Cys Ala Cys His Tyr Asn Cys Pro Gly Asp Asn Asp Ile Phe Glu
325 330 335

Ser Leu Tyr Tyr Arg Lys Met Tyr Gly
340 345

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: ~~SEQ ID NO:5:~~

CGCGGGATCC TCGCGACAC AATGAGCT

28

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGCGGGTACC AGGTAGCATT TAGTCCCTAA

30

SUB
B

668040-51284660

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAAGGATCCA CAATGAGCTC CCGAATC

27

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCTCTAGAT TAAGCGTAGT CTGGGACGTC GTATGGGTAT TGGAACAGCC TGTAGAAG

58

SUB
B1

668040-57887650